**Introduction**

Climate, being a sample statistic of weather, is an important phenomenon that affects the whole world every year. According to the definition of the World Meteorological Organization (WMO), climate is average weather over a longer period. The stability of climate parameters for a variable of interest is an important criteria because any instability may result in unwanted consequences. To clarify this discussion, we refer here to the unnatural rainfall pattern of the monsoon in the Indian subcontinent during July to September 2010 [1]. The rainstorm structures compiled by Tropical Rainfall Measuring Mission (TRMM) satellites also explained that the pattern of rainfall during 2010 was unnatural [2].

Principally, changes in the rainfall pattern of Pakistan and the China-Tibetan Plateau region was linked to low pressure in the meandering of the subtropical jet stream [3]. The whole period of flooding was physically connected with the disastrous natural event of the Russian heat wave-wildfire termed as a “blocking

---

**Abstract**

Flooding is a serious and devastating hazard that many countries face regularly. Timely detection of these changes can help us manage rainwater and hence flooding. Numerous statistical tools are available for the early detection of these changes. In this regard, application of control charts is an effective choice toward monitoring natural events. In this study, we used self-proposed and existing control charts as a means of early detection of changes in the rainfall data of Pakistan as a case study. The proposed methodology covered two aspects: one to enhance the ability of Shewhart control charts toward detection of small or moderate changes in the behavior of natural events, and another to offer skewness correction based on individual control chart under runs rules for managing these natural events – especially when collected data follows unknown skewed distribution. Results elucidate that control charts structures were efficient in early detection and prediction of floods during 2010. Our results are in accordance with the theoretical results of existing studies, and we propose similar methods to be used for meteorological purposes.

**Keywords**: control chart, CUSUM, early warning, EWMA, floods, rainfall, runs rules, non-normality
event” in the jet stream. As a result, rainfall intensity in Pakistan and northwestern India was amplified due to more fuel recharge (moisture content), and caused one-fifth of Pakistan to be inundated (cf. [3]). However, unnatural changes in parameters of climate variable may have been detected and predicted in a timely way in order to avoid any hazards.

Monitoring of natural or unnatural changes in the parameters of interest (i.e., location or dispersion) can be achieved using statistical process control (SPC). The SPC consists of different tools such as histogram, preto diagram, cause and effect diagram, and control charts. Among different SPC tools, practitioners and researchers generally prefer a quality control chart due to an efficient methodology and wide applications. For instance, the application of control charts can be seen in analytical measurements [4], medical sciences [5], industrial sciences [6], agricultural sciences [7], and environmental sciences [8, 9]. Likewise, applications of control charts can be seen for monitoring cyclone impact, fire frequency and fisheries management [10].

The quality control charts can be classified into memory-less (also termed as Shewhart-type) and memory types. To exemplify this classification, Shewhart-type control charts used in the study of [7, 8, 11, 12] represent the memory-less control chart, whereas CUSUM and EWMA control charts used in the study of [4, 9, 10] indicate a memory control chart.

The Shewhart-type individual control chart based on one point decision rule under the assumption of normal probability distribution [8] is generally considered efficient for detection of large changes in the parameter of a variable of interest and less efficient for the detection of moderate or small changes. This issue can be resolved by incorporating more runs rules or sensitizing rules. For more details on the illustrated issue, along with a remedy in the form of runs rules, can be seen in the literature of quality control charts such as [6, 7, 12, 13]. On the other hand, memory control charts have been proven to be efficient for the detection of small or moderate changes and less efficient for detecting large changes. For readers concern, we define the one-point decision rule in the contacts of a climate variable as follows: a climate event is said to be out-of-control if any measured observation of the variable of interest breaches outside the baseline or in-control limits.

In this study, we propose a normality-based individual control chart under runs rules for monitoring the location parameter and then make an attempt to implement annual behavior of rainfall data of Pakistan. The purpose of implementation is to show the significance of attaching more runs rules in terms of diagnosing small or moderate changes in a natural phenomenon (e.g., rainfall), which often is the reason for floods in many countries. This statement has been taken as a first subject of the current study. Afterwards, the second subject of the study is to propose skewness correction based on individual control chart structure incorporated with runs rules. The skewness correction-based structure depends on the amount of skewness of the collected data instead of the restricted assumption of normal probability distribution. The details of skewness correction methodology can be seen in the study of [13, 14].

Besides, we implement a cumulative sum (CUSUM) control chart and an exponentially weighted moving average (EWMA) control chart on the annual behavior of rainfall data in order to demonstrate their importance for early detection of disasters. This has been considered as a third subject of study. Moreover, we have chosen those optimum control charts from the list of memory-less and memory control charts whose application results may verify the theoretical results of the study of [3, 15]. Another reason for the selection of these control charts is the motivation from the study of [8-10]. The present study provides detailed insights of our earlier study [16], which was later encouraged by [17] in the monography titled “identification of overlapping features in time series data”.

**Experimental**

Daily rainfall data of all provinces are collected from the Pakistan Meteorological Department (PMD), which is located at Flood Forecasting Division, Lahore, Pakistan. This rainfall data represents the behavior of 10 consecutive monsoon periods from May 2008 to September 2010, among which different rainfall stations were tested using a variety of control charts. On the same guideline of [8], rainfall data from January 2000 to March 2008 is selected for a defining baseline limit (in-control limit), while the last three consecutive monsoon periods are plotted and monitored through control charts. Mean rainfall rate of each month is calculated by taking an average of daily measured rainfall (in mm).

Now we provide necessary details of proposed normality-based individual control charts under runs rules, proposed skewness correction-based individual control chart under runs rules, a cumulative sum (CUSUM) control chart and exponentially weighted moving average (EWMA) control chart by taking into account subjects of the current articles.

**Proposed Normality-Based Individual Control Chart under Runs Rules**

In order to achieve the first subject, runs rules of the study of [6] is defined in the aspect of rainfall as a variable interest: A rainfall level for a certain time period can be considered as out-of-control if at least \( r \) observations out-of \( m \) consecutive observations breaches outside the control limit. Now we have incorporated these runs rules with the usual design structure of an individual control chart (upper sided) as follows:
Application of Quality Control Charts...

\[ UCL = \mu + Z_{(r,m,1-p)} \sigma \]  

(1)

...where \( \mu \) and \( \sigma \) are known mean and standard deviation of the rainfall for a particular area respectively; \( Z_{(r,m,1-p)} \) is the control limit multiplier of standard normal distribution, which depends on a given value of favorable observations \( r \) in order to declare out-of-control, total observations \( m \) under considerations and probability of single observation \( p \) of rainfall going outside the control limit \( (UCL-) \). The value of \( p \) should be picked in such a way that the false alarm rate \( \alpha \) of an independent runs rule remains stable at a desired level [6]. The false alarm rate \( \alpha \) indicating the predefined probability that at least \( r \) out of \( m \) consecutive observations of rainfall go out-of-control when actually the rainfall process is assumed to be in-control. Moreover, an appropriate value of \( p \) can be picked by solving the following expression [6]:

\[ \alpha = \sum_{r=1}^{m} \left( \frac{m!}{(m-r)!r!} \right) p^r (1-p)^{m-r} \]  

(2)

The control chart structure given in Equation 1 remains valid for the case of known parameters. In order to cover the case of unknown parameters, Equation 1 can be written as:

\[ UCL = \bar{X} + Z_{(r,m,1-p)} \frac{MR}{1.128} \]  

(3)

...where \( \bar{X} \) is the average of collected rainfall data and \( MR \) is the average of moving range for rainfall data of the particular area. In order to accomplish the first subject of the current study, we will implement two runs rules (1/1 and 2/3) with \( \alpha = \frac{1}{ARL_0} = 0.05 \) (see Table 1) on the annual rainfall data. The specific choice of in-control average runs length \( ARL_0 \) is motivated by the study of [9]. On similar lines, one can make an attempt for the other choice of runs rules, \( \alpha \) and \( ARL_0 \).

Proposed Skewness Correction-Based Individual Control Chart under Runs Rules

The structure given in Equation 1 is helpful for monitoring the natural events when collected data on a variable of interest follows normal distribution. In order to cover the case of departure from normality such as skewed distribution, we have made an attempt to provide a skewness correction-based individual control chart structure under runs rules [16] as follows:

\[ UCL = \mu + \left( Z_{(r,m,1-p)} + C^*_i \right) \sigma \]  

(4)

For the case of unknown parameters, the structure (4) is presented as:

\[ UCL = \bar{X} + \left( Z_{(r,m,1-p)} + C^*_i \right) \frac{MR}{d^*_2} \]  

(5)

...where \( d^*_2 \) is the skewness correction-based unbiasing constant [14] that depends on level of skewness \( k_3 \) of collected data instead of the assumption of normal distribution. Furthermore, \( C^*_i \) is the skewness adjustment factor that depends on \( Z_{(r,m,1-p)} \) and the amount of skewness \( k_3 \). The expression for computing the \( C^*_i \) is as follows:

\[ C^*_i = \frac{1}{6} \left( Z^2_{(r,m,1-p)} - 1 \right) k_3 (1 + 0.2 k_3^2) \]

The rest of the quantities, such as \( \bar{X} \), \( Z_{(r,m,1-p)} \) and \( MR \), were already defined earlier. The procedure for constructing the control limits (Equation 5) followed by plotting the rainfall data is given in the following steps: step-1: compute average \( (\bar{X}) \), average of moving range \( (MR) \), skewness \( (k_3) \) and pick the adjustment factor \( (C^*_i) \) of collected rainfall data (see Table 2); step-2: input all required quantities (obtained from steps 1-2) into Equation 5; step-3: after constructing the control limits in step-2, plot the individual observations of collected rainfall data through control charts (note: Equations 3 and 5 serve similar purposes, and even one can expect similar outcomes; the only difference is that for implementing the proposed normality-based control chart structure we have to test the assumption of normality of collected rainfall data, whereas for implementing the skewness correction-based individual control chart we have to just compute the skewness of data).

Cumulative Sum (CUSUM) Control Chart

CUSUM control chart is also called memory control chart because it possesses the ability to attach current measured value with the successive values of the past. The upper side charting statistic of the CUSUM control chart, denoted as \( C^*_i \), can be defined as follows:

Table 1. \( Z_{(r,m,1-p)} \) for varying choices of \( \alpha \).

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>0.010</th>
<th>0.020</th>
<th>0.030</th>
<th>0.040</th>
<th>0.050</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule (1/1)</td>
<td>2.33</td>
<td>2.053</td>
<td>1.88</td>
<td>1.75</td>
<td>1.64</td>
</tr>
<tr>
<td>Rule (2/3)</td>
<td>1.58</td>
<td>1.41</td>
<td>1.30</td>
<td>2.29</td>
<td>1.16</td>
</tr>
</tbody>
</table>

Table 2. \( C^*_i \) for specific choices of \( \alpha \) and \( k_3 = 0.77 \).

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>0.010</th>
<th>0.020</th>
<th>0.030</th>
<th>0.040</th>
<th>0.050</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule (1/1)</td>
<td>0.4920</td>
<td>0.3588</td>
<td>0.2829</td>
<td>0.2302</td>
<td>0.1902024</td>
</tr>
<tr>
<td>Rule (2/3)</td>
<td>0.1670</td>
<td>0.1086</td>
<td>0.0766</td>
<td>0.4717</td>
<td>0.0373</td>
</tr>
</tbody>
</table>
\( C_t^+ = 0 \)
\( C_t^- = \max\{0, X_t - \bar{X} - k + C_{t-1}\} \), \( t = 1, 2, 3 \ldots m. \)

In the above expression, \( X_i \) denotes the \( ith \) observation of measured rainfall, \( \bar{X} \) represents average rainfall for a particular area, and \( k \) is the reference value that is approximately picked half of the shift that needs to be detected in the rainfall behavior. Now the upper-sided CUSUM control chart structure is given below:

\[
H = h \frac{MR}{1.128}
\]  

(6)

The procedure for plotting statistic \( C_t^+ \) is very simple, that is calculate each time the value of \( C_t^+ \) and plot it against the control limit \( H \) for each rainfall observation. Any single value outside the signaling limit (\( H \)) is declared as out-of-control. Moreover, the value of \( h \) depends on the value of \( ARL_0 \) and \( k \). In this study, we will show the application of the CUSUM control chart for monitoring the rainfall data by considering the value of \( ARL_0 = 20 \) and \( k = 25 \) (Table 3). Also, the choice of \( ARL_0 \) and \( k \) for monitoring the environmental variable has motivated from the study of [9]. For a required choice of \( ARL_0 \) and \( k \), a different value of \( h \) can be seen in the study of [18].

### Exponentially Weighted Moving Average (EWMA) Control Chart

The idea of an exponentially weighted moving average (EWMA) control chart can be exploited for an efficient and quick detection of small shifts in process parameters. Similar to the CUSUM control chart, the EWMA control chart is based on the entire sequence of rainfall data instead of just using the current information like the Shewhart-type control chart. The chart uses a varying weight scheme; assigning highest weight to the most recent observation and the weights decrease exponentially as observations get less recent. For the observed rainfall \( (X_t) \) at time \( t \), the EWMA statistic \( (W_t) \) is given as:

\[
W_t = 2X_t + (1 - \lambda)W_{t-1}, \quad t = 1, 2, 3, \ldots m, \text{ and the control chart structure is as follows:}
\]

\[
UCL = \bar{X} + L \frac{MR}{1.128} \sqrt{\frac{\lambda}{2 - \lambda}} (1 - (1 - \lambda)^{2t})
\]  

(7)

Table 3. Value of \( h \) and \( k \) for certain choices of \( ARL_{l_0} \).

<table>
<thead>
<tr>
<th>( ARL_{l_0} )</th>
<th>200</th>
<th>150</th>
<th>100</th>
<th>50</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h )</td>
<td>5.60</td>
<td>5.09</td>
<td>4.41</td>
<td>3.33</td>
<td>2.113</td>
</tr>
<tr>
<td>( k )</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Table 4. Values of \( L \) and \( \lambda \) for different choices of \( ARL_{l_0} \).

<table>
<thead>
<tr>
<th>( ARL_{l_0} )</th>
<th>200</th>
<th>150</th>
<th>100</th>
<th>50</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>( L )</td>
<td>2.405</td>
<td>2.285</td>
<td>2.099</td>
<td>1.75</td>
<td>1.20</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
</tr>
</tbody>
</table>

\( L(0 < \lambda < 1) \) is the weight assigned to the most recent rainfall observation, and \( L \) is the control limit coefficient depending on the values of \( \lambda \) and \( ARL_{l_0} \). The EWMA control chart triggers an out-of-control signal for any value of \( W \) lying outside the above control limits. The choice of \( \lambda \) creates flexibility in the EWMA charting structure for practitioners. If the detection of small shifts in the process parameters is of major concern (as in the case of rainfall data), small values of \( \lambda \) (say less then 0.30) are most useful. Otherwise, if the detection of large shifts is desirable, large values of \( \lambda \) are mostly recommended. Note that the proposed individual control chart under runs rules (described in Equation 1) became a special case of EWMA chart when \( \lambda = 1 \), \( r = 1 \) and \( m = 1 \). Similar to the CUSUM control chart, the choices of \( ARL_{l_0} \) and \( \lambda \) for showing the application EWMA control chart are taken as 20 and 0.25 (see Table 4) respectively. Moreover, different values of \( L \) for a given value of \( ARL_{l_0} \) and \( \lambda \) can be seen in the study of [4].

### Results and Discussion

#### Application of Control Charts

In this section we intend to show the application of the above-mentioned control charts in climate. Before implementing the control charts on rainfall data, we applied the log transformation on the collected rainfall data in order to fulfill the assumption of normality. The data transformation mechanism can also be seen in other studies such as [8] and [7]. Mehmood et al. [7] used the log-transformation on electric conductivity of land in order to transform the data into the normal distribution. It is worth mentioning that for the case of the proposed skewness correction-based control chart under runs rules we have calculated the skewness of the collected rainfall data instead of employing the mechanism of log-transformation. This clearly indicates an advantage of the skewness correction-based control chart over a normality-based control chart in the form of utilizing original data instead of making certain transformations.

Gove et al. [8] emphasize one important thing that should be taken into consideration while constructing a process control chart – especially when dealing with monitoring of climate variables: the control limit for declaring out-of-control climate change. A common strategy for defining control limit (baseline limit) is to...
use the early portion of in-control data as a baseline period [10]. Here, the procedure of constructing the control limit (based on the early portion of in-control data) is termed as retrospective or Phase-1 analysis [19]. The control limits obtained from the retrospective analysis are further used for early detection of signals of upcoming (future) flooding hazards. Hence, data from January 2000 to March 2008 are used for constructing an in-control limit (Equations 3, 5-7), while the last three consecutive monsoon periods are plotted and monitored through control charts.

Illustration of Results, Discussion and Comparisons

We have implemented an individual control chart under runs rules, CUSUM control chart and EWMA control chart as shown in Figs 1-4. These charts are representing rainfall behavior of Sargodha city, which was selected due to appropriate geography, and presenting nature of data. Herby, time interval between vertical red lines is illustrating the monsoon period in the Indian Subcontinent. Now we illustrate results of each of the control charts and also give brief comparisons among each other.

Fig. 1. Normality-based individual control chart for Sargodha (Punjab).

Fig. 2. Skewness correction-based individual control chart for Sargodha (Punjab).

Fig. 3. CUSUM control chart for Sargodha (Punjab).

Fig. 4. EWMA control chart for Sargodha (Punjab).
Proposed Control Charts under Runs Rules

Proposed individual control charts structures under runs rules (Equations 3, 5) are elucidating the rainfall behavior of the last three monsoon periods from June 2008 to August 2010 for Sargodha (Punjab), which is approximately at 597 feet elevation from mean sea level (Fig. 5). Out-of-control signals detected during the monsoon period are showing the natural or usual phenomenon (Figs 1-2). In Figs 1-2, two out-of-control signals can be seen during the monsoon periods of 2008 and 2010 respectively. However, a large shift is detected during 2010, which was very severe compared to previous monsoon periods. This is sufficient evidence to declare that these control charts are in accordance with theoretical interpretation of [3]. Furthermore, Rule (2/3) has detected smaller changes in the pre-monsoon period of 2010, which was not observed in 2008 and 2009. Actually, shifts in the pre-monsoon period of 2010 clearly indicate early warning signals of an upcoming flooding hazard. This early detection ability of control charts for floods seems to be advantageous compared to other statistical techniques. Thus, the following results clearly indicate that attaching more runs rules with the design structure of individual control charts enhances their ability toward the detection of early warning signals. Finally, we have met the first and second subjects of the current article (see Section 1).

Cumulative Sum (CUSUM) Control Chart

In Fig. 3, CUSUM control chart is applied to demonstrate the behavior of rainfall for the same area mentioned above (i.e., Sargodha). Results illustrate that the CUSUM control chart do not detect abrupt larger shifts that happened during the monsoon period of 2009. However, upward shifts during April 2010 are indicators of upcoming intense rainfall, and the floods occur in early August. This clearly shows that performance of the CUSUM control chart for monitoring rainfall behavior can be considered as more reliable for small or moderate shifts.

Exponentially Weighted Moving Average (EWMA) Control Chart

The EWMA control chart is applied (see Fig. 4), whose results revealed that it is not good in detecting abrupt large shifts that happened during the monsoon period of 2009. This renders the performance of an EWMA control chart similar to the CUSUM control chart in the case of a sudden shift. Furthermore, the role of an EWMA control chart and proposed individual control chart under runs rules seems similar while detecting small shifts, i.e., early warning signals. Nevertheless, both control charts are giving early warning signals of intense rainfall in April 2010 and the floods occur in early August. Thus, control chart results are in accordance with the published fact of flooding by [15]. Similarly, early warning signals detected by the CUSUM control chart can also be seen in April 2010.

Comparative Analysis of Control Charts

Based on the illustrated results, we found that proposed individual control charts under runs rules...
(i.e., 2/3), CUSUM control chart, and EWMA control chart are more proficient in detecting early warning signals compared to individual control charts based on the one-point decision rule (i.e., 1/1). However, in the case of larger shifts, the Shewhart individual control chart based on one-point decision rule is performing better as compared to other control charts under consideration. In addition, the Shewhart control charts with one point decision rule has declared rainfall out-of-control when a sudden or abrupt change (natural or unnatural) occurred, whereas the Shewhart control chart with other decision rules (i.e., 2/3) has not claimed the process to be out-of-control because it relies on 2 out-of-control observations from 3 observations. Here it is practically cleared that the one-point decision rule should be preferred over other sensitizing rules when a sudden or large change occurred in a variable of interest (e.g., rainfall). For instance, two recent studies have been seen in the literature of statistical process control on the importance and implementation of runs rules [20, 21]. On the same guidelines, we constructed control charts for rainfall data of other provinces whose results were also consistent to the trends of intense rainfall and flooding during the monsoon period of 2010.

Conclusions

Each climate variable is important in the natural hazard system and the changes in these variables may result in a natural disaster. Hence, monitoring these variables is of prime importance and should be done in a timely way. To approach this idea, we showed the application of different control charts including proposed normality-based individual control charts under runs rules, proposed skewness correction-based individual control chart under runs rules, CUSUM control chart and EWMA control chart on the rainfall behavior of Pakistan.

Results indicated that control charts can be a useful tool in the detection of large, smaller, and moderate shifts in natural hazard studies. Among different control charts, a proposed control chart under extra sensitizing rules (i.e., 2/3), CUSUM and EWMA are found to be more helpful for the early prediction of flooding hazard, whereas the usual individual control chart based on the one-point decision rule (i.e., 1/1) performed outstandingly when a sudden abrupt change occurred in the rainfall for a particular time period. The advantages of these statistical techniques may help policymakers manage such unusual events before they result in calamities. Moreover, our results have been validated by comparing them with theoretical results of existing studies.

Conflict of Interest

The authors declare no conflict of interest.

References